DTIC FILE COPY

AFHRL-TP-90-40

AIR FORCE

AD-A226 743

HUMAN RESOURCES

PERCEPTUAL INSPECTION TIME: AN EXPLORATION OF TACTICS TO ELIMINATE THE APPARENT-MOVEMENT STRATEGY

Scott R. Chaiken

MANPOWER AND PERSONNEL DIVISION
BROOKS AIR FORCE BASE, TEXAS 78235-5601

Robert K. Young

Department of Psychology
University of Texas
Austin, Texas 78713

September 1990
Interim Technical Paper for Period February 1989 - April 1990

LABORATORY

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235-5601

)

NOTICE

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Public Affairs Office has received this paper, and it is releasable tot he National Technical Information Service, where it will be available to the general public, including foreign nationals.

This paper has been reviewed and is approved for publication.

WILLIAM E. ALLEY, Technical Director
Manpower and Personnel Division

ROGER W. ALFORD, Lt Col, USAF
Deputy Chief, Manpower and Personnel Division

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 1990	Interim Paper - February 1989 to April 1990

4. TITLE AND SUBTITLE
Perceptual Inspection Time: An Exploration of Tactics to Eliminate the Apparent-Movement Strategy

5. FUNDING NUMBERS
PE - 61102F
PR - 2313
TA - T1
WU - 33

6. AUTHOR(S)
Scott R. Chaiken
Robert K. Young

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Department of Psychology
University of Texas
Austin, Texas 78713

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAMES(S) AND ADDRESS(ES)
Manpower and Personnel Division
Air Force Human Resources Laboratory
Brooks Air Force Base, Texas 78235-5601

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFHRL-TP-90-40

11. SUPPLEMENTARY NOTES
This paper is based on collaborative research performed during the second author's residence at AFHRL/MOEL, Brooks AFB, Texas, under a grant from the USAF Summer Faculty Research Program through Universal Energy Systems, Inc., SFRP-GSRP Office.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)
The research reported herein is of broad relevance to the construction of culture-fair intelligence tests based upon processing-speed measurements. Three experiments used an inspection-time (IT) paradigm with a post-test questionnaire to segregate subjects into an apparent-movement strategy group and a no-strategy group. Experiment 1 replicated previous research by finding a significant IT/intelligence-test correlation only for subjects from the no-strategy group. Experiment 2 attempted to influence the probability of adopting the strategy by altering the standard IT task parameters. Only increased strategy use could be produced by our manipulations. Experiment 3 replicated Experiment 1 but added an alternative IT task designed to eliminate the strategy. The alternate IT task predicted IQ less well than for the no-strategy group in the standard task, suggesting the alternate was also susceptible to an unknown set of strategies. Finally, across all experiments, no reliable differences were found for the two groups of subjects with respect to IT or intelligence test performance.

14. SUBJECT TERMS
computerized testing
individual differences
inspection time
intelligence testing

15. NUMBER OF PAGES
25

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

AFHRL Technical Paper 90-40 September 1990

PERCEPTUAL INSPECTION TIME: AN EXPLORATION OF TACTICS TO ELIMINATE THE APPARENT-MOVEMENT STRATEGY

Scott R. Chaiken

MANPOWER AND PERSONNEL DIVISION
BROOKS AIR FORCE BASE, TEXAS 78235-5601

Robert K. Young
Department of Psychology
University of Texas
Austin, Texas 78713

Reviewed and submitted for publication by

Patrick C. Kyllonen, Research Scientist
Learning Abilities Research Function

SUMMARY

The research reported herein is of broad relevance to the construction of culture-fair intelligence tests based upon processing-speed measurements. Three experiments used an inspection-time (IT) paradigm with a post-test questionnaire to partition subjects into an apparent-movement strategy group and a no-strategy group. Experiment 1 replicated previous research by finding a significant IT/intelligence test correlation only for subjects of the no-strategy type. Experiment 2 attempted to influence the probability of adopting the strategy by altering the standard (i.e. Experiment 1's) IT-task parameters. Only increased strategy use could be produced by our manipulations. Experiment 3 replicated Experiment 1 but added an alternate, non-line length-discrimination IT task designed to eliminate the strategy. The alternate IT task predicted IQ less well than for the no-strategy group in the standard task, suggesting the alternate was also susceptible to an unknown set of strategies. Finally, across all experiments, no reliable differences were found for the two types of subjects with respect to mean performance on the standard IT or intelligence-test performance.

TABLE OF CONTENTS

		Page
I.	INTRODUCTION	1
II.	EXPERIMENT 1	2
	Method	2
	Subjects	2
	Apparatus	2
	Procedure	2
	Inspection Time Task	3
	Verbal Analogies Task	3
	Number-Series Test	3
	Results	3
	Discussion	5
III.	INTRODUCTION TO EXPERIMENT 2	5
	Method	5
	Subjects	5
	Apparatus	5
	Procedure	5
	Results	6
	Strategy Use and Validities	6
	Main Effects	7
	Comparisons Between Experiments	7
	Discussion	8
IV.	INTRODUCTION TO EXPERIMENT 3	8
	Method	9
	Subjects	9
	Apparatus	9
	Procedure	9
	Inspection-Time Tasks	10
	Results	10
	Discussion	12
V.	CONCLUSIONS	12

REFERENCES 14

APPENDIX: INSPECTION-TIME QUESTIONNAIRE 17

LIST OF TABLES

Table **Page**

1 Intercorrelations and Reliabilities (Diagonals) for the No-Strategy and Strategy Groups of Experiment 1 4

2 Intercorrelations and Reliabilities (Diagonals) for the No-Strategy and Strategy Groups of Experiment 3 11

PERCEPTUAL INSPECTION TIME: AN EXPLORATION OF TACTICS TO ELIMINATE THE APPARENT-MOVEMENT STRATEGY

I. INTRODUCTION

The present paper re-examines an experimental task and its associated phenomena that have attracted wide attention in the intelligence literature. The task is called inspection time (IT) and is defined as the minimum presentation time necessary to make a reliable judgment between two or more stimuli. This paradigm was first mentioned in 1886 by Cattell (cited in Deary, 1986). The modern-day inspection-time (IT) paradigm was first developed by Vickers, Nettelbeck and Willson in 1972 but attracted little interest until a 1976 paper by Nettelbeck and Lally reported a correlation of .92 between IT performance and intelligence-test performance (hereafter, IQ). Since then many papers have been published dealing with the IT/IQ relationship (cf. Lubin & Fernandez, 1986).

The usual IT task relies on visual discrimination although auditory versions of the task have also been employed (e.g. Raz & Willerman, 1985). The IT stimulus consists of two thin vertical lines connected at the top by a thin horizontal line. One of the two vertical lines is shorter than the other, and the subject's task is to tell which leg is shortest. An IT trial consists of a warning signal, a test-stimulus, and finally a backward mask consisting of vertical equal-length solid bars that cover the informative parts of the IT stimulus. Task difficulty is manipulated by varying the stimulus onset asynchrony (SOA) of the backward mask which terminates the presentation of the IT stimulus. The SOA is the interval of time beginning with the onset of the IT stimulus and ending with the onset of the backward mask. In general, the SOA is varied in an adaptive manner until the subject is responding at some high-level of accuracy (e.g. 97%). The subject's IT is then defined as the SOA at which the subject has reached the accuracy criterion (Vickers, Nettelbeck & Willson, 1972). In the present studies we employ a paradigm in which the IT SOA is varied across several fixed levels of stimulus duration while the subject's accuracy is observed at each presentation duration. This non-adaptive procedure has also been employed in several different IT studies (e.g. Anderson, 1986; Wilson & Nettelbeck, 1986).

While the IT paradigm is straightforward in what it tries to measure, it is not without a major methodological problem. The backward mask creates apparent movement (Egan, 1986) such that an auxiliary cue other than a length difference can be used in making the discrimination judgment. Typically, IT performance correlates with IQ in the 30s or 40s (Pearson's $\underline{r}$s); however, when subjects who use apparent movement to improve their scores are separated from those who do not, the IT/IQ correlation drops to non-significance. The IT/IQ relation becomes more significant for those who don't use a motion strategy (Mackenzie & Bingham 1985; Mackenzie & Cumming, 1986). Though this work may be criticized for using small sample sizes, the replication of the effect on two separate occasions is intriguing. Mackenzie et al. (1985, 1986) employed sample sizes of 37 and 29; furthermore, the no-strategy group numbered less than half the sample. Even with such small sample sizes, correlations for the no-strategy group were observed in the 60s and 70s.

At issue in the current report are problems dealing with the use of motion strategies. In particular, it is not decidely known whether the use of the motion strategy is something a subject can elect to do or is the result of some deeper difference between the users and non-users. Initial evidence is mixed. Evidence that strategy use is not elected comes from a study which shows that the apparent movement cues can not be taught to the non-users and therefore is not a strategy per se (Mackenzie et al., 1985). However, the same laboratory has also reported that users and non-users are not different in the minimum stimulus asynchrony necessary to perceive apparent motion between two dots appearing in close temporal and spatial succession on a dark background (Mackenzie et al., 1986).

Experiment 1 is a replication of the IT/IQ correlation and its interaction with strategy use using a relatively large sample from a homogeneous population of normal subjects (Air-Force basic trainees). Once a replication is demonstrated, we can explore the issue of whether the validity of IT to IQ can be boosted by changing the parameters of the IT task in Experiment 2. This contrasts with direct instructional interventions as was done in Mackenzie et al., (1985). Finally, the logic of changing the parameters of an IT task is taken further in Experiment 3, which is an exploration of an alternate IT task where motion strategies are plausibly absent.

II. EXPERIMENT 1

Method

Subjects

Subjects were 113 Air-Force basic trainees in their 6th day of basic training (ages 17 to 27; 87 males, 26 females). All subjects had completed high school (or passed an equivalence exam, the CLEP) and some had varying levels of college experience.

The subject-ability range was restricted owing to Air-Force policies of selectively recruiting applicants above a certain percentile rank on the Armed Forces Qualifying Test (AFQT) a composite of scores from the Armed Services Vocational Aptitude Battery (ASVAB). The ASVAB is a broad measure of crystalized intelligence (Ree, Mullins, Mathews, & Massey, 1982). In our sample, the minimum percentile score observed was 38, and the mean AFQT percentile rank was 63.3 with a standard deviation of 15.33. The norming population on which the percentile score is based has a mean rank of 50.7 and a standard deviation of 29.13.

Apparatus

Zenith 248 computers with EGA color monitors and standard keyboards were used to present stimuli and collect data.

Procedure

After initial instructions each subject was given 144 trials on the IT task followed by a 14-item questionnaire dealing with strategies, perceived attributes of the IT task, and perceived attributes of the subject's performance. Only one question (question 7) pertained specifically to the use or non-use of movement cues as a strategy that improved the subject's score. Following the questionnaire, Scale 3 of the Cattell Culture Fair Test of *g* was administered (Cattell & Cattell, 1959). This is a speeded test administered in about 20 minutes and is not intentionally verbal or quantitative but involves figural reasoning to induce the nature of progression in a sequence of abstract patterns. Subjects used test booklets to encode specific questions but all instructions and responses were computer-managed (constituting a change in the standard paper-and-pencil administration procedure).

Also administered by computer, as auxiliary measures of intelligence, were tests of verbal analogies taken from Sternberg and Nigro (1980) and a Number-series Test taken from Tirre (1988). These auxiliary intelligence tests are described in a later section.

The subjects were run in groups of about 24. Each subject was tested in an individual testing carrel.

Inspection Time Task

Every trial on the IT task included the presentation of a warning signal ("*" for 500 msecs., followed by 200 msecs of blank-screen), the test stimulus (for a variable duration), and the backward mask (for 500 msecs.), after which the subject's response was recorded. A new trial was given after a 500 msecs. delay.

The test stimuli in the IT task was composed of a short vertical line and a long vertical line (10mm and 30mm, respectively), that were 5mm apart. The width of the lines (and their top horizontal connector) was two pixels. The mask had two equal-length legs (50mm in length) which started at the horizontal connector at the same points as the test-stimulus legs. Each leg of the mask was two mm wide and completely covered the thinner vertical lines of the test stimulus. Both the test stimulus and the masks were the same color (yellow).

Eight different stimulus to mask SOAs were employed (17, 33, 50, 67, 83, 100, 117, and 133 msecs). Within successive blocks of 16 trials, each of the eight presentation durations was used twice, once with the shorter leg of the test stimulus on the left and once with it on the right. All conditions were presented in a different random order within each block for each subject.

After presentation of the test stimulus the subject indicated on which side the shorter line appeared by pressing the "L" or "D" keys for the right and left sides, respectively. Knowledge of results was given after each of the first 16 trials to be sure the subject understood the instructions. Subsequently, the subjects were told how many times they were correct only after the end of each 48-trial period.

Verbal Analogies Task

The 60 Sternberg and Nigro test items are verbal-reasoning analogy problems. They were preceded by detailed instructions which demonstrated and explained the formats of questions and the general task demands of finding parallel relationships between pairs of English words. The formats were as follows: (1) A : B :: C : ?, (2) A : B :: ? : ?, (3) A : ? :: ? : ?, where A, B, C represent the "given" part of each analogy, and question marks represent the information that subjects had to induce. Subjects determined which alternative, among several presented underneath the problem statement, was the best and entered their response by typing the number of that alternative. The task was self-paced and produced a percent correct score.

Number-Series Test

The Number-series Test (Tirre, 1988) is patterned after Thurstone's (1935) Number-series Test which displays a sequence of numbers with one position empty. The subject needs to induce what number should go into the empty position from a list of 4 alternatives presented underneath the sequence. Subjects responded by typing in the number (1 through 4) of the alternative they thought correct. These 26 problems had sequences that ranged from five to six numbers and could be completed by noticing simple (ascending or descending) progressions, by noticing exponential progressions in the differences between items, and by noticing the existence of two interleaved (i.e. alternating) sequences.

Results

The number of correct responses, collapsed over SOAs, was the principle predictor used for assessing the IT/IQ relation. (A d-prime description of a subject's line-length discrimination ability and a minimum SOA necessary to achieve a 75% accuracy criterion was also considered as different forms of the IT predictor. These measures yielded very similar results to those reported below, and so only one measure is reported.)

A significant correlation between the Cattell test and the IT task was observed ($r(111) = .34$; $p < .01$) but not for the other intelligence tests and IT ($r(111) = .13$ and $r(111) = .17$ for verbal-analogies and number-series tests, respectively).

Next, the group was divided into those individuals who reported using apparent movement to improve their scores and those subjects who reported not using such a cue. The correlation matrix in Table 1 reflects both groups, with the no-strategy group in the upper matrix and the strategy group in the lower matrix. The 36 subjects who reported that they did not use a movement cue showed higher relationships between IT and different measures of IQ. Furthermore, the correlation of IT to the average of the (z-scored) intelligence measures was found to be significantly larger for the no-strategy group ($z = 2.2$; $p < .05$). Hence, the modulating effect of apparent-movement has been replicated in our population using our experimental procedures.

Table 1. **Intercorrelations and Reliabilities (Diagonals) for the No-Strategy and Strategy Groups of Experiment 1**

No-Strategy group		I	II	III	IV
I	Cattell	.65			
II	Verbal Analogy	.30	.93		
III	Number Series	.49*	.41*	.72	
IV	Inspection Time	.56**	.37	.32	.95
N of cases: 36					
Strategy group					
I	Cattell	.46			
II	Verbal Analogy	.37**	.88		
III	Number Series	.22	.28*	.39	
IV	Inspection Time	.21	.01	.09	.92
N of cases: 77					

*$p < .01$.
**$p < .001$.

Some general analyses were performed to investigate possible differences between the strategy and no-strategy groups that may have been responsible for the validity differences found. The two groups did not differ on mean scores for IT or for any of the intelligence tests although the no-strategy group did exhibit a larger variance with respect to the Cattell test ($F(35,76) = 2.31$; $p < .002$) and for the number-series test ($F(35,76) = 1.77$; $p < .04$).

The two groups showed no difference in overall IT efficiency, and no difference in their improvement in IT performance as SOA increased, although the full sample showed reliable linear and quadratic effects ($F(1,111) = 248$; $p < .001$; $F(1,111) = 176$; $p < .001$) associated with SOA. There was a practice effect for IT across blocks of trials (performance went from 80% to 85% correct, $F(8,888) = 3.90$; $p < .001$), but no interaction of practice with strategy use.

Discussion

The results of the present experiment replicate previous findings (c.f. Lubin & Fernandez, 1986) in that a significant correlation between IT and intelligence-test scores was found. The use of movement cues destroyed the IT/IQ relationship (c.f. Egan, 1986), with such use being present in about two-thirds of the sample. However, the IT/IQ correlation for the no-strategy group sometimes exceeded the correlations among the IQ tests.

Initial attempts to find differences between the movement-users and non-users were unsuccessful, except for an indication that non-users exhibit a broader range of test scores on the Cattell Test, and possibly the number-series test.

Finally, if the use of a movement cue is, in fact, a strategy, it is probably not a good strategy at all levels of IQ. High IQ subjects apparently reduce their IT performance when using movement cues but low IQ subjects apparently enhance their performance. At least this is consistent with the finding of no mean differences in IT performance between the strategy and no-strategy groups even while their IT validity to IQ appears to be different.

III. INTRODUCTION TO EXPERIMENT 2

If the use of movement cues could be eliminated, performance on the IT task might be a much better predictor of intelligence. This speculation assumes that the movement-cue users who are prevented from using the cue would perform similarly to the non-users. Thus the emphasis in Experiment 2 was in comparing conditions that might maximally foster and maximally hinder the use of apparent movement as a cue. The hypothesis is that an IT task that hinders the use of a movement cue will show a lower frequency of movement-cue users and a stronger IT/IQ relationship relative to an IT task that maximally fosters such cues. Such should be the case if use of the motion strategy is determined principally by task characteristics and not person characteristics.

Method

Subjects

Subjects were 120 Air Force basic trainees in their 6th day of basic training (all male) drawn from the same population as in Experiment 1. This sample had a mean AFQT percentile rank of 64.3 with a standard deviation of 15.5.

Apparatus

The same equipment and testing carrels were employed in this study as employed in Experiment 1.

Procedure

Experiment 2 differed from the "standard" IT paradigm (i.e. that paradigm used in Experiment 1) in many respects. The most salient differences were the masks used to prevent further processing of the IT stimuli and the complexity of the IT task itself.

In an easy condition, the task parameters were changed so that the use of the movement cue could be expected to increase. This was accomplished by making the mask very similar to the test stimulus which should increase the impression that the masks are continuations of the lines of the IT stimulus. Line masks, the same length as the solid-bar masks of Experiment 1, were used. These masks simply extended the IT stimulus lines to be the same length.

In contrast, the parameters employed in a hard condition were meant to decrease the use of movement cues. In the hard condition a "meta-contrast" mask (Kahneman, 1968) such that an unfilled rectangle (the same dimensions as the bar masks used in Experiment 1) surrounded and replaced the test stimulus after the appropriate presentation duration. In addition to being physically dissimilar from the lines they cover up, the metacontrast type of mask masks at a deeper level of perceptual analysis than light-burst masks which mask at a more peripheral (i.e. retina-bound) level of analysis (Haber & Hershenson, 1980). The light-burst mask may encompass the solid-bar mask of Experiment 1 and the the line-extension mask of the easy condition; hence, assuming the apparent movement strategy reflects deeper perceptual analyses, the meta-contrast mask could be a more potent mask.

Two other aspects of the hard condition also differ from the standard paradigm. First, the vertical lines that made up the IT stimulus had length ratios of two to three rather than a one to three ratio employed in the easy condition and in Experiment 1. The intent of this modification was to reduce the salience of the movement cue by reducing the magnitude of the apparent movement in the hard condition. Finally, the hard condition used a same/different rather than a forced-choice response paradigm. This meant that the subject could receive two short lines or two long lines as a test stimulus requiring a same judgment, or two lines that differed in length, requiring a different judgment. For the hard condition, subjects responded "L" for same length and "D" for different lengths. We also expected this difference (from the standard IT task and the easy condition) to decrease the use of motion cues, because the subject would have to compare two movements in contrast to simply deciding where the movement had occurred.

There were also many surface changes in the IT task from Experiment 1 to 2. These surface changes were present in both the easy and hard conditions. Mask length was reduced to 46mm and the length of the long line was increased to 36mm (making the short-line 12mm for the easy condition and 24mm for the hard condition). Finally, the IT stimulus lines were moved from 5mm to 25mm apart, the warning signal was extended to 2 seconds, and the two longest SOAs from Experiment 1 were dropped.

Half the subjects of a testing session were randomly assigned to the easy condition, with the remaining half assigned to the hard condition. Subjects received 6 blocks of 24 trials, with each duration occurring 4 times per block. Each of the two possible stimuli in the easy condition and each of the four possible stimuli in the hard condition appeared an equal number of times at each duration within a block. The order of conditions was random across blocks and subjects. Subjects received the same questionnaire as used in Experiment 1 for both the easy and hard conditions, after which they received Scale 3 of the Cattell Culture-Fair Test, administered as in Experiment 1.

Results

Strategy Use and Validities

The principal results of this experiment involve two separate issues: differences between conditions in the use of movement cues, and differences between conditions in the IT/IQ relationship. Recall that our intent was to devise two IT tasks which differed in their susceptibility to the motion strategy. The easy IT task was expected to increase the use of movement cues while the hard was expected to decrease such use.

A CHI-square test was applied to the two conditions of Experiment 2 to judge whether the frequency of movement-users differed according to whether the condition was easy or hard;

the chi-square was not significant, there being nine non-users in the easy condition and ten non-users in the hard condition.

With regard to validity, neither easy nor hard condition showed any significant positive relation to IQ. Correlations were $r(58) = .12$ and $r(58) = -.01$ respectively. When the IT performance for each condition was *z*-scored within that condition and the correlation of the transformed scores to IQ was assessed for all non-users, again no validity was found ($r(17) = .10$). One possible reason for a lack of validity is a lack of reliability in both the easy and hard conditions of the IT task. However, reliabilities for the conditions of Experiment 2, estimated via the split/half method, compared favorably with Experiment 1 (reliabilities were .93, .95, and .85 for Experiment 1, the easy, and the hard condition, respectively).

Main Effects

The main effect of condition in Experiment 2 was highly significant ($F(1,118) = 98.3$; $p < .001$), the hard condition having a mean of 62% correct and the easy condition having a mean of 84% correct. There was also a significant interaction of duration by condition ($F(5,590) = 16.9$; $p < .001$). Essentially, there was less improvement in accuracy for the hard condition compared to the easy condition as stimulus-to-mask SOA was increased. Accuracy in the hard condition went from chance 52% to 71% correct as stimulus duration increased from minimum to maximum; while accuracy in the easy condition went from 63% to 90% correct. Also, accuracy asymptoted for the easy condition at around 90% at the 67 msec SOA, but accuracy increased linearly with SOA for the hard condition.

Comparisons Between Experiments

Another set of analyses compared the conditions of Experiment 2 to the more standard conditions of Experiment 1. A CHI-square test compared the use of movement cues across Experiment 1 and Experiment 2. This test was significant (X (2) = 8.33; $p < .016$). The movement users accounted for 68%, 85%, and 83% for Experiment 1, the easy condition, and the hard conditions respectively.

A MANOVA compared the easy condition of Experiment 2 to the more standard condition of Experiment 1. Only the common SOAs between the two experiments were used in this analysis. No main effects or interactions with duration were found (all F's greater than .5 but less than 1). Hence, the "easy" condition is not easier than the more standard IT task, even while showing a difference in the prevalence of the strategy. However, when another MANOVA was run that included the additional post-hoc variable of strategy use, a significant strategy-use by experiment interaction was observed, $F(1,169) = 10.3$; $p < .002$. Given the great disparity of cell n for this last MANOVA analysis (minimum sample size of 9, maximum sample size of 77), we decided to explore the robustness of this interaction using focused non-parametric tests.

Recall that no difference was observed between the no-strategy and strategy-groups of Experiment 1. In contrast, a Mann-Whitney test comparing the 51 strategy users and the 9 non-users for the easy condition of Experiment 2, revealed a significant rank difference ($U = 110$; $z = -2.46$; $p < .0132$, two-tailed). Non-users had a mean rank of 17.2 while users had a mean rank of 32.8. Hence, being a non-user in a condition which plausibly fosters the use of movement cues was related to poorer performance on the easy IT task.

Another Mann-Whitney test compared the strategy users from Experiment 1 to the strategy users of the easy condition of Experiment 2. This comparison was also significant (mean ranks were 55.9 and 77.5, respectively, $U = 1302.5$; $z = -3.22$; $p < .0013$, two-tailed). Hence, for people who elect to use the strategy, the task parameters of the easy condition are easier than the task parameters of Experiment 1.

Finally, a Mann-Whitney test comparing the users to the non-users in the hard condition was not significant (mean ranks 31.2 and 27.0, respectively; $\underline{U}$ = 215; $\underline{z}$ = -.68). Hence, strategy use was irrelevant to the performance in the hard condition which plausibly diminishes the usefulness of the strategy, but not its attractiveness to the subject.

Discussion

While there was no apparent difference between the frequencies of non-users in the hard and easy conditions of Experiment 2, there was a difference in strategy-use frequencies when Experiment 2 and Experiment 1 are compared. Both easy and hard conditions have more movement-cue users than expected on the basis of Experiment 1's observed frequencies. Why the hard condition increased the use of movement (or more precisely the tendency to say movement was used to improve ones score) is unclear. On the other hand, use of the movement strategy appeared to improve performance in the easy condition.

The effect of a shift in the frequency of the apparent movement strategy is important because it indicates maleability of strategy use in the IT task. Prior to this experiment the only other attempts to modify strategy use in the IT task have been negative (Mackenzie et al., 1985). Additionally, the lack of validity in conditions which show a relative preponderance of strategy use (i.e. Experiment 2 as a whole) is not surprising. As in previous research, the IT/IQ relation exists only in subjects that do not use the motion strategy.

IV. INTRODUCTION TO EXPERIMENT 3

Because the manipulation of mask and test-stimulus parameters did not work entirely as expected, we decided to employ a different strategy for finding an IT task free of the influence of movement cues. Instead of working with the standard IT paradigm, an alternative task was devised in which quantity discriminations were made by the subject. The subject had to determine whether 2, 3, or 4 Xs were displayed (randomly) on a grid. The target Xs became masked after a variable SOA by a grid completely filled with Xs. This task was thought to be so different from the usual IT task that movement use would not be a factor. The hypothesis is that the *x*-counting task will be more valid against IQ measures than the more standard IT paradigm, owing to the non-existence of a movement strategy.

A further issue in Experiment 3 involves how the strategy and no-strategy groups differ. Experiment 1 indicates no difference between the two groups in terms of their mean IT and IQ performance; however, a distributional difference between the two groups can also account for the difference in validities observed. For instance, the difference could arise from a greater variance of measured IQ within the no-strategy group.

Several findings from Experiment 1 suggest a distributional difference may be responsible for the validity differences. The variance of the Cattell test scores was significantly greater for the no-strategy group, and it was the Cattell test that showed the largest validity difference for the two groups. Also the no-strategy group tended to show systematically higher reliabilities on intelligence measures which is consistent with a broader range of intelligence levels in that group. Ignoring the issue of why the questionnaire could have partitioned the population of subjects into high-variance and restricted-variance groups, one can easily test this explanation for the validity differences.

If one collects a reasonably-sized sample for the no-strategy group, then according to a different-distributions theory, even a motion-irrelevant predictor of intelligence should have its validity affected by the strategy/no-strategy classification. For instance, the verbal-analogies test would certainly be considered a motion-irrelevant predictor of the Cattell Culture Fair Test and vice versa. And if the no-strategy group had greater variance in the underlying *g*-construct, one would expect verbal analogies to predict Culture Fair better for that group. In fact, for Experiment 1, two out of the three intercorrelations between intelligence tests were higher in the no-strategy group.

If, on the other hand, the classification of strategy versus no-strategy is completely local to the IT task, then no differential validity for a motion-irrelevant predictor of intelligence should be observed. Hence, in Experiment 3 we include the auxiliary intelligence tests given in Experiment 1, to see if the trends described above will replicate with a larger sample of no-strategy subjects.

Method

Subjects

A total of 261 subjects from the same population as the previous experiments was employed. There were 185 males and 76 females. The mean percentile AFQT score was 68.5 with a standard deviation of 16.4.

Apparatus

The same apparatus and testing set-up were employed as in the previous experiments.

Procedure

The experimental session was composed of five blocks of testing that involved IT, intelligence, and reaction-time tests.

The first three blocks consisted of an IT task, followed by Scale 2 and Scale 3 of the Cattell Culture-Fair Test, followed by another IT task. Scale 2 of the Culture-Fair Test was not included in Experiments 1 and 2 because it was unavailable. That scale is easier than Scale 3 and is generally administered if the obtained scores on Scale 3 are very low. Half the subjects received the *x*-counting task, the Cattell tests, and the more standard IT task in that order, while the other half received the reverse ordering. For the Cattell tests Scale 2 was always given before Scale 3. Both the *x*-counting task and IT task had similar questionnaires with one question involving the use of a movement cues (the standard IT questionnaire was left unchanged while the *x*-counting questionnaire was slightly different to accommodate the changed task). After the Cattell and IT tasks, subjects received a 5-minute break and then the second half of the testing session.

Following the break came the auxiliary intelligence tests which were interleaved with some simple reaction-time tests. The intelligence tests employed were Sternberg and Nigro's (1980) verbal-analogies test as used in Experiment 1 and a revised version of the Number-series Test (Tirre, 1988). The revisions on the latter were mainly in the task's instructions but extra problems were also constructed to improve reliability (Tirre, 1988). Subjects received 24 five-position (four-filled) problems and 12 eight-position (seven-filled) problems. The test problems were preceded by five examples which explained and demonstrated the kinds of problems the subject would encounter. The examples were meant to provide a mini-tutorial on how to accomplish the task of number series inductions.

The reaction-time tests were a speeded physical-identity task and a speeded semantic-verification task. Intelligence tests and reaction-time tests were randomly assigned to their position in the testing sequence with the constraint that intelligence tests alternate with reaction-time tests with intelligence-tests leading.

The reaction-time tests were included principally as explorations. Since no remarkable correlations between IT and RT occurred, and no *unexpected correlations between RT and IQ* measures occurred, the RT tasks are not detailed further.

Total testing time ranged from 1.5 to 2.0 hours.

Inspection-Time Tasks

The administration of the standard IT task was the same as in Experiment 1 except for the addition of six more blocks of testing (96 more trials) and the test-ordering differences noted above.

The *x*-counting task began with computer-presented instructions and six example trials that explained and demonstrated the task at slower durations than the subject received during the *course of the task* (and subjects were made aware of that last fact). This practice block was geared to introduce subjects to the variety of stimuli they would receive. The practice items consisted of 2 replications each of 2x, 3x, and 4x trials. Each subject received the same practice items.

The *x*-counting task consisted of 216 trials that came in 12 blocks of 18 trials. Each block of 18 trials included every combination of stimuli (2, 3, or 4 xs) with SOAs from 16 to 100 msecs in six equal steps. Subjects received trials within block without interruption, but were allowed to rest and receive accuracy feedback at block boundaries.

The trial sequence was as follows. First a warning signal (a "*" presented at the center of the 7 x 15 grid) was given for 450 msecs, followed by a blanking of the screen for another 100 msecs. The test stimulus was then presented and replaced by the backward mask immediately after the appropriate duration lapsed. After 500 msecs the screen blanked and the subject responded by pressing either the "2," "3," or "4" key at the top of the keyboard. If the subject did not respond after a 3-second interval, a prompt would appear that told the subject, "The computer is waiting for your response (2, 3, or 4). Subjects received feedback after every trial (the words "correct" or "incorrect" appeared on the center screen).

The stimulus items were generated randomly with the constraint that the xs could not occur in the middle (3 by 3) portion of the grid. This constraint was imposed because pilot studies indicated that the warning signal was sometimes confused for an x. Every subject saw the same stimuli but not in the same order.

Results

Table 2 presents the correlation matrix for the no-strategy group (upper matrix) and the strategy group (lower matrix) for all relevant measures employed in the study. About 8% of the sample was deleted owing to list-wise deletion; however, the correlations are not appreciably different when only pair-wise correlations are considered. The basic patterns of correlations from Experiment 1 have been replicated, with the Cattell and verbal analogies test showing the largest differential correlations for users and non-users of the motion strategy.

Table 2 provides evidence against the idea that users and non-users of the motion strategy have different distributions of intelligence levels. The ability of one intelligence test to predict another does not differ across groups as it should if a differing breadth of intelligence levels is selected out by the strategy/no-strategy distinction. Reliabilities for intelligence-test scores are also more equivalent across groups compared to Experiment 1.

Table 2. **Intercorrelations and Reliabilities (Diagonals) for the No-Strategy and Strategy Groups of Experiment 3**

No-Strategy group		I	II	III	IV	V
I	Cattell	.75				
II	Verbal Analogy	.56**	.86			
III	Number Series	.57**	.54**	.86		
IV	Inspection Time	.44**	.29*	.21	.88	
V	X Account Time	.30*	.24	.26*	.39**	.95

N of cases: 88

Strategy group		I	II	III	IV	V
I	Cattell	.62				
II	Verbal Analogy	.58**	.83			
III	Number Series	.55**	.51**	.82		
IV	Inspection Time	.19*	.14	.18	.92	
V	X Account Time	.27**	.12	.11	.23*	.94

N of cases: 154

Note. Strategy is defined with respect to the standard IT task.
*$p<.01$.
**$p<.001$.

Other negative evidence comes from direct comparisons of the distributions of intelligence-test scores for the strategy and no-strategy groups. The variance of intelligence-test scores did not differ significantly between the two groups for any of the intelligence measures used.

As in Experiment 1, a set of analyses were performed to assess the effects of stimulus duration on accuracy for the IT task. Again no main effect of strategy use was found. Although there was a significant SOA effect ($F(7,1799) = 1182$; $p<.001$), there was no interaction of stimulus duration with strategy use. Thus, as in Experiment 1, the presence of the strategy did not differentially help or hurt the strategy group.

With regard to the *x*-counting task, we were disappointed in its overall validity to intelligence-test scores (computed as the average of the *z*-scored intelligence tests) as compared to a more standard IT task. The two validities were virtually identical ($r(240)$ =. 27; $p<.001$, and $r(240) = .28$; $p<.001$ for *x*-counting and IT tasks, respectively). The finding of equal validities goes against the hypothesis that an IT task that plausibly lacks the movement strategy should lead to higher validities. Interestingly both sorts of IT tasks made independent contributions to the prediction of intelligence for the complete sample (partial correlations .20 and .21 for the *x*-counting and IT tasks, respectively; multiple $R = .34$).

The low validity of the *x*-counting task (relative to expectations) can not be attributed to the presence of a motion strategy. The frequency of non-users (as assessed by a parallel questionnaire) was substantially larger in the *x*-counting task than in the more standard IT task (i.e. 60% versus 36%), and unlike the more standard IT task, the non-users in the *x*-counting task showed no differential validity from those that claimed they used movement in that task.

We should note that we would have predicted motion strategies to be non-existent in the *x*-counting task. Even so, most of the strategy fluctuation in going from the IT task to the *x*-counting task was owing to the strategy group shifting over to a no-strategy classification (as expected). One fifth of the non-users for the IT task claimed to switch over to a motion strategy in the *x*-counting task; however, half of the strategy-users from the IT task switched over to a no-strategy classification in the *x*-counting task.

Discussion

Experiment 3 was a more radical attempt to remove the salience of movement from the IT paradigm. Unfortunately, the (largely) motion-irrelevant *x*-counting task is not as valid an IT task as the standard IT task is for the no-strategy group, suggesting considerable room for improvement in designing an IT task that purifies the "g"-related construct IT measures.

Perhaps some unknown strategy is present in the *x*-counting task that reduces its validity. This notion is plausible given the correlations between the two different types of IT task. While the correlation is largest among the no-strategy group (suggesting the construct measured by IT is "purer" in that group), the correlation between any pair of intelligence tests is much higher than the correlation between the two IT tasks. This occurs in spite of the higher reliabilities shown by the IT tasks.

Finally, Experiment 3 fails to replicate the trends indicating a distributional difference in IQ levels sampled for the strategy and no-strategy groups. No variance differences among any of the IQ measures were observed and there was no difference in how well an intelligence test predicted another intelligence test as a function of group. However, like Experiment 1, the effect of strategy use was detectable only in its effect on the validities of IT to IQ (and then only for the standard task). Again, no differences were observed in the discrimination ability for the two groups at any stimulus duration.

V. CONCLUSIONS

These experiments have replicated the findings of small-n studies showing that the use of a movement cue decreases the validity of an IT task as a predictor of intelligence-test scores. Substantial samples of subjects were obtained that reported not using the apparent-movement strategy, making possible, for the first time, detailed inspections of differences in the speed of information extraction in the IT task. No main effects of strategy use were found for IT conditions which replicate validity differences between the two types of subject. This suggests that the strategy under investigation is not a typical strategy that boosts performance at all levels of IQ. Instead it must be hurting performance for as many people as it helps. There were also no distributional differences in IQ observed for the strategy and no-strategy groups for an experiment large enough to make distributional comparisons meaningful. Hence, identifying a person as not using the strategy does not have an effect on validity through some bimodal filtering mechanism that only lets those high in IQ and those low in IQ into the no-strategy group.

Experiment 2, demonstrates that one can make the apparent-movement strategy both more attractive and more effective for the subject by altering parameters of the IT task. This contrasts with previous findings that the strategy is not teachable. We think there is still a lot of work needed in the area of experimental manipulation. In particular, the practical goal of finding conditions that reduce the apparent-movement strategy is a necessity before any theoretical research about the IT/IQ relationship can be accomplished. Our research suggests so far that there is only a very narrow window of task parameters that make the standard type of IT task (i.e. line-length discriminations) strategy free. We infer this from the finding that the IT/IQ relationship decreases and the reported use of movement cues increases with our manipulation of task parameters.

Experiment 3, demonstrates how serious the strategy problem is in general for IT research. Here we found evidence that the reliable variance observed in two dissimilar IT tasks was mainly task-specific. If this result generalizes to other diverse IT tasks then perhaps the only way to develop a strategy-free IT task is to sample performance from many diverse IT tasks.

The diverse-task way of purifying the IT construct may be viable if such assessment could be shown to be free of learning and transfer effects among IT tasks. Such effects would confound perceptual speed with some general perceptual learning ability. To the best of our knowledge, transfer effects between dissimilar IT tasks is minimal, since there were no order effects for the standard IT and *x*-counting tasks in Experiment 3. However, these two tasks only correlate weakly; hence, the challenge in this approach is to find diverse tasks that both correlate moderately well and afford little transfer. This may well be as challenging as trying to find a single strategy-free IT task.

REFERENCES

Anderson, M. (1986). Inspection time and IQ in young children. *Personality and Individual Differences*, 7, 677-686.

Cattell, R.B., & Cattell, A.K.S. (1959). *Handbook for the Individual or Group Culture-Fair Intelligence Test: Scale 3*. Champaign, IL: IPAT.

Deary, I.J. (1986). Inspection time: Discovery or rediscovery? *Personality and Individual Differences*, 7, 625-631.

Egan V. (1986). Intelligence and inspection time: Do high-IQ subjects use cognitive strategies? *Personality and Individual Differences*, 7, 695-700.

Haber, R.N., & Hershenson, M. (1980). *The psychology of visual perception*. New York: Holt, Rinehart, and Winston.

Kahneman, D. (1968). Method, findings, and theory in studies of visual masking. *Psychological Bulletin*, 70, 404-425.

Lubin, M.P., & Fernandez, J.M. (1986). The relationship between psychometric intelligence and inspection time. *Personality and Individual Differences*, 7, 653-657.

Mackenzie, B., & Bingham, E. (1985). IQ, inspection time, and response strategies in a university population. *Australian Journal of Psychology*, 37, 257-268.

Mackenzie, B., & Cumming, S. (1986). How fragile is the relationship between inspection time and intelligence: The effects of apparent-motion cues and previous experience. *Personality and Individual Differences*, 7, 721-729.

Nettelbeck, T., & Lally, M. (1976). Inspection time and measured intelligence. *British Journal of Psychology*, 67, 17-22.

Raz, N., & Willerman, L. (1985). Aptitude-related differences in auditory information processing: Effects of selective attention and tone duration. *Personality and Individual Differences*, 6, 299-304.

Ree, M.J., Mullins, C.J., Mathews, J.J., & Massey, R.H. (1982). *Armed Services Vocational Aptitude Battery: Item and factor analyses of Forms 8, 9, and 10* (AFHRL-TR-81-55, AD-A113 465). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Sternberg, R.J., & Nigro, G. (1980). Developmental patterns in the solution of verbal analogies. *Child Development*, 51, 27-38.

Thurstone, L.L. (1935). Primary mental abilities. *Psychometric Monograph*, (No. 1). Chicago: University of Chicago Press.

Tirre, W.T. (1988). *Number-series test*. Unpublished experimental test.

Vickers, D., Nettelbeck, T., & Wilson, R.J. (1972). Perceptual indices of performance: The measurement of "inspection time" and "noise" in the visual system. *Perception*, 1, 263-295.

Wilson, C., & Nettelbeck, T. (1986). Inspection time and the mental age deviation hypothesis. *Personality and Individual Differences*, 7, 669-675.

APPENDIX: INSPECTION-TIME QUESTIONNAIRE

Note: All questions remained on the screen while the subject's answer (numeral and content) was displayed with the further question: Is this correct? Yes (press y) or No (press n). If the subject typed n, the question and follow-up check would repeat. The subject was not allowed to continue to the next question until y was typed. The questions were given in the order presented below. Question number 7, response alternative 4, marked the person as not using a motion strategy.

1. This task was
   1 Easy
   2 Sometimes easy, sometimes hard
   3 Hard

2. Were you comfortable doing this task?
   1 Very uncomfortable
   2 Uncomfortable
   3 Neither uncomfortable or comfortable
   4 Comfortable
   5 Very comfortable

3. How often did you have to guess?
   1 Not very often
   2 Occasionally
   3 Often
   4 Very often

4. Did you improve over time?
   1 Yes
   2 No

5. Did you see any of the following?
   1 Movement of either or both lines
   2 Movement of the second part covering the lines
   3 Both 1 and 2
   4 Neither 1 or 2

6. Did you see any of the following?
   1 Blinking or flashing of the lines
   2 Fading or "washing out" of the lines
   3 Both 1 and 2
   4 Neither 1 or 2

7. Did any of the following help you on this task?
   1 Movement of the lines
   2 Movement of the second part
   3 Both 1 and 2
   4 Neither 1 or 2

8. Did any of the following help you on this task?
   1 Not focusing on the *
   2 Washing out the lines
   3 Both 1 and 2
   4 Neither 1 or 2

9. When you found the task to be difficult what BEST describes the difficulty?
   1 The short line was easy to see but finding its position was difficult
   2 The short line seemed to be the same length as the long line
   3 The lines were not seen at all before the part covering them came on
   4 The task was very easy except for occasional random mistakes

10. Were the lines presented too quickly to see?
   1 Rarely
   2 Sometimes
   3 Often
   4 Very often

11. Was the task easier if the short line was on the
   1 Left
   2 Right
   3 No difference

12. Did you try to improve your score in any of these ways?
   1 Blinking your eyes
   2 Looking for movement
   3 Both 1 and 2
   4 Neither 1 or 2

13. Did you use any of these methods to help your score?
   1 Looking for the lines to blink
   2 Looking for the lines to fade in or out
   3 Both 1 and 2
   4 Neither 1 or 2

14. Did you use some method to improve your score?
   1 I used one or more of the methods listed previously
   2 I used a method not listed previously
   3 Both 1 and 2
   4 I did not use any method